Research Article

Weighted Complex Network Analysis of Shanghai Rail Transit System

Yingying Xing, 1 Jian Lu, 2 and Shendi Chen 3

1 School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai, China
2 School of Transportation Engineering, Tongji University, 4800 Cao’an Road, Shanghai, China
3 Shanghai Maritime University, College of Transport and Communication, Shanghai, China

Correspondence should be addressed to Shendi Chen; sdchen@shmtu.edu.cn

Received 30 March 2016; Revised 13 June 2016; Accepted 10 July 2016

1. Introduction

There is a rapidly growing literature on the complex networks present in transport systems and complex network analysis is a useful method to analyze the structure of transport systems. For example, studies of worldwide airport networks have shown small-world property and exhibited heavy tailed power-law distributions [1]. For the public transportation systems in Poland, various network properties, such as the distribution of degree and clustering coefficient, have been analyzed [2]. Further, the national highway network of Pakistan has been investigated with weighted complex network analysis of travel routes on the network [3]. Rail transit systems are in essence physical networks that are composed of stations or stops all linked by rails. Latora and Marchiori [4] found that Boston public transportation system exhibits the small-world behavior. Angeloudis and Fisk [5] studied the world’s largest subway systems and found that systems with substantial shared track are less robust than dedicated line systems of similar size. Lee et al. [6] analyzed statistical properties and topological consequences of the Seoul subway system and found that the flow weight distribution exhibited a power-law behavior. Soh et al. [7] contributed a complex weighted network analysis of travel routes on the Singapore rail and bus transportation systems. Zhang et al. [8] summarized the universal characteristics of the urban rail transit networks. Besides topological characteristics of networks, the reliability and robustness of metro networks were also widely studied. By looking at 33 metro systems in the world, Derrible and Kennedy [9] analyzed the complexity and robustness of metro systems and provided insights/recommendations for increasing the robustness of metro networks. Based on complex network theory, Zhang et al. [10] studied the connectivity, robustness, and reliability of the Shanghai subway network. De-Los-Santos et al. [11] provided passenger robustness measures for a rail transit network.

Nevertheless, these literatures, from a complex network perspective, focus more on topological features than dynamic traffic flow. The quantity of traffic in large transport
### 2. Weighted Networks Data

Shanghai RTS network consists of 286 nodes denoting stations and 317 edges accounting for a link connecting two nearest stations. The average degree of the network is \(\langle k \rangle = \frac{2E}{N} = 2.22\), while the maximal degree is 8. As already observed in previous literatures [8, 10], the topology of the network exhibits both scale-free and small-world properties. Datasets that are provided by the Shanghai Shentong Metro Company list the hourly in and out passenger flows for each RTS station and passenger flows between adjacent stations. In this study, passenger flows during morning peak hours from 7:00 to 9:00 in a typical weekday are analyzed, during which the highest volume on a weekday could be observed.

Due to the method of data capture, it is important to note that the paper is not only investigating the topological and functional properties of the RTS network, but also the passenger flows between the different stations. In this work, it is assumed that typical travel was bidirectional, and hence the weight \(w_{ij}\) of one edge between a pair of nodes (stations) \(i\) and \(j\) is defined to be the sum of passenger flows in both directions and \(w_{ij} = w_{ji}\).

### 3. Weighted Network Analysis of Shanghai Rail Transit Network

In this section, we present a topological and dynamical analysis of the Shanghai RTS network. Table 1 shows comparison analysis of basic network properties of unweighted network (topological) and weighted network (dynamical). As mentioned above, the Shanghai RTS network is comprised of 286 nodes and 317 edges and the average degree of the network is 2.22. The clustering coefficients of the RTS network (including the unweighted network and the weighted network) are 0.0012 and 0.0024, respectively, which both indicate that the local connectivity of the RTS network is very poor. The characteristic path length \(L\) of the unweighted network between two nodes \(v_i\) and \(v_j\) is defined in terms of the network distance, which represents the minimum number \(n_{ij}\) of links necessary to go from node \(v_i\) to node \(v_j\). The network diameter refers to the maximum shortest path and the diameter \(d = 41\) for the unweighted network. The network efficiency describes the global connectivity of the network and is given by

\[
E = \frac{2}{N(N-1)} \sum_{i<j} \frac{1}{n_{ij}}. \tag{1}
\]

But it is more appropriate to use the physical distance \(d_{ij}\) rather than the network distance \(n_{ij}\) in measuring the real network efficiency \(E\). In terms of the physical distance, the characteristic path length, diameter, and network efficiency for the weighted network are given by \(L = 24.49\ km, D = 116.20\ km,\) and \(E = 0.0352\), respectively.

#### 3.1. Degree and Strength Distribution

In a topological network, for a given node \(i\), its degree is the number of edges shared with other nodes and defined as \(k_i = \sum_{j} a_{ij}\). In a weighted network, a more meaningful measure of the network properties is obtained by introducing strength \(s_i\), defined as

\[
s_i = \sum_{j=1}^{N} a_{ij} w_{ij}. \tag{2}
\]

This quantity \(s_i\) combines node degree with edge weight and reflects centrality of a node \(i\) in the weighted network.

The probability distribution of \(p(s)\) exhibits heavy tailed behavior and shows similarities with the degree distribution \(p(k)\) (see Figure 1). Meanwhile, the strengths of stations appear scale-free (indicating the existence of hub nodes with very high traffic) and follow a power-law distribution

\[
p(s) \sim s^{-\gamma}, \tag{3}
\]

with \(\gamma \approx 0.487\).

When each edge in the network is assigned a weight, researchers naturally would like to know the correlation between weight and topology structure of the network. In order to shed more light on the relation between the node strength and degree, the dependence of the strength \(s_i\) on degree \(k_i\) is investigated. As it can be seen in Figure 2, the distribution of the average strength \(s(k)\) as function of the degree \(k\) of nodes can be well approximated by the power-law behavior

\[
s(k) \sim k^\theta, \tag{4}
\]

with an exponent \(\theta \approx 1.1\). This reveals that the strength of nodes is positively associated with the degree and coincides with the fact that the more the connections a station is
linked to, the more the traffic it handles. This observation also implies that the rail transit traffic grows much faster than the number of connections. But it should also be noted that in the Shanghai RTS network the largest degree node is the Century Avenue, while the highest strength node is the People’s Square (see Figure 3). What is more, although many nodes share similar degree, the traffic handled by each rail station may differ significantly. In consequence, the dynamical properties of a network may differ significantly from its topological properties. To summarize, the node strength and degree indicate the importance or connectivity of node in the network from the different angles, and the node strength is more appropriate to describe the real RTS network. In addition, the strength of nodes during peak hours provides a measure for the characteristic capacity of the facilities near a station.

3.2. Topological and Dynamical Clustering. The clustering coefficient is a measure of local cohesiveness and connectivity of the network. It is defined for any node $i$ as the fraction of connected neighbors and can be expressed by the following equation:

$$ C_i = \frac{2E_i}{k_i(k_i-1)} = \frac{2}{k_i(k_i-1)} \sum_{j,h} a_{ij}a_{ih}a_{jh}, \quad (5) $$

where $E_i$ refers to the number of edges between the neighbor nodes of node $i$. The average clustering coefficient $C_i$ is averaged over clustering coefficient of all nodes in the network and can be expressed by

$$ \langle C \rangle = \frac{\sum_i^N C_i}{N}. \quad (6) $$

It is important to note that $C_i$ is defined solely on topological grounds. Edge weights and their correlations may change our view of the hierarchical and structural organization of the
4 Discrete Dynamics in Nature and Society

network (see Figure 4). In order to solve the previous incongruities, Barrat et al. [1] introduce the weighted clustering coefficient $C_w$ which takes into account the weights of edges:

$$C_w(i) = \frac{1}{s_i (k_i - 1)} \sum_{j<k} \frac{w_{ij} + w_{ik}}{2} a_{ik} a_{jk} a_{ki}. \quad (7)$$

Similarly, the average weighted clustering coefficient is given by

$$\langle C_w \rangle = \frac{\sum_{i=1}^{N} C_w(i)}{N}. \quad (8)$$

If the weights are completely uncorrelated in the network, $C_w = C$ and $C_w(k) = C(k)$ [13]. However, weights in real-world networks are often correlated, leading to two possible situations. When $\langle C_w \rangle$ is larger than $\langle C \rangle$, triangles in the network are more likely connected by edges with larger weights. Conversely, when $\langle C_w \rangle$ is smaller than $\langle C \rangle$, the topological clustering is generated by edges with smaller weights.

In this case, the weighted clustering coefficient of the Shanghai RTS network is $\langle C_w \rangle = 0.0024$, approximately twice as large as their topological measures (see Table 1). As $\langle C_w \rangle$ is larger than $\langle C \rangle$, it can be concluded that triangles are more likely constructed by edges with larger weights. However, although Shanghai RTS network features the longest mileage in the world, $\langle C \rangle$ and $\langle C_w \rangle$ of the Shanghai RTS network are significantly smaller than other international metropolises such as Tokyo, New York, and London [8]. This result suggests that connectivity of the network should also be considered when the RTS networks are planned and designed.

3.3. Degree-Degree Correlation. Degree-degree correlation demonstrates the extent of a node’s degree related to the average degree of its neighbors and reflects the node’s connection preference [14]. If high-degree nodes in the network tend to link with each other, the network is considered to be assortative. On the contrary, if high-degree nodes tend to connect with low-degree nodes, the network is a disassortative network. The average nearest neighbors degree $k_{nn}$ is one of the common indexes in measuring degree-degree correlation and is defined as

$$k_{nn} = \frac{1}{N} \sum_{j \in N} a_{ij} k_j. \quad (9)$$

Further information can be gathered by inspecting the function $k_{nn}(k)$; it represents the average degree of all $k$-degree nodes $N_k$ (neighbors of all nodes with $k$-degree) and can be expressed as

$$k_{nn}(k) = \frac{1}{Np(k)} \sum_{i \in N_k} k_{nn,i}. \quad (10)$$

If $k_{nn}(k)$ is an increasing function of $k$, the network is assortative. If $k_{nn}(k)$ is a decreasing function of $k$, the network is disassortative.

In weighted networks, the weighted average nearest neighbors degree is defined by

$$k_{nn}^w = \sum_{j \in N} w_{ij} a_{ij} k_j = \frac{1}{S} \sum_{j \in N} a_{ij} w_{ij} k_j. \quad (11)$$

This quantity measures the affinity of a node to connect to high or low-degree neighbors. This depends on the edge weights between itself and neighbors. If $k_{nn}^w$ is larger than $k_{nn}$, the highly weighted edges connect to the large-degree neighbors. Conversely, the edges with the smaller weights are pointing to the neighbors with lower degree when $k_{nn}^w < k_{nn}$.

Returning to the Shanghai RTS network, Figure 5 shows that both the weighted network and the unweighted network are definitely assortative at least up till $k = 6$. It is surprising to find that both $k_{nn}^w(k)$ and $k_{nn}(k)$ decrease at $k = 8$. After conducting a deep research, we find that the network has one and only one node with degree $k = 8$. Therefore, from the overall trend, the Shanghai RTS network can be considered...
to be assortative. This finding is in line with findings in other studies mentioned above [1]. Also, since \( k_{nn,j} > k_{nn,j} \) in the whole \( k \) spectrum, the heavily weighted edges are linked to the large-degree neighbors; in other words, edges with the larger flow of passengers pass through more well-connected locations.

3.4. Topological and Dynamical Rich-Club Coefficient. Rich-club coefficient (RCC) is an important quantity and introduced as a measure of the interconnectivity between hub nodes. For the set of \( N_{r,k} \) rich nodes whose degree is larger than \( k \), the RCC is defined as

\[
\phi(k) = \frac{2E_{\geq k}}{N_{\geq k}(N_{\geq k} - 1)},
\]

where \( E_{\geq k} \) represents the number of edges between \( N_{r,k} \) rich nodes in the club. In other words the RCC \( \phi(k) \) measures whether or not connections are established among rich nodes. If there are no edges between rich nodes, then \( \phi(k) = 0 \), whereas \( \phi(k) \) reaches the value of 1 when all the possible edges are present. However, it is easy to find that, even without any correlation, one edge is more likely to be shared between two rich nodes rather than two low-degree nodes. In particular, Colizza et al. [15] find that in an uncorrelated graph the RCC increases as \( k^2 \) for large \( k \). In order to overcome this problem, they compared \( \phi(k) \) measured on the real network with the corresponding \( \phi_{null}(k) \) obtained from an appropriate null model and proposed the ratio

\[
\rho(k) = \frac{\phi(k)}{\phi_{null}(k)}.
\]

Similarly, Opsahl et al. [12] extended the definition of \( \phi(k) \) and \( \rho(k) \) for weighted networks. They made a generalization by defining a rich-club coefficient for the weighted case as

\[
\phi^w(r) = \frac{E_{\geq r}}{\sum_{r=1}^{E_{\geq r}} w_n^{rank}},
\]

where \( r \) is a richness parameter, such as strength. \( E_{\geq r} \) is the sum of weights between nodes whose richness is larger than \( r \) and \( E_{\geq r} \) is the number of nodes whose richness is larger than \( r \). \( w_n^{rank} \) is the \( n \)th weight in \( E_{\geq r} \) strongest links within the whole network. For example, Figure 6 shows a schematic representation of a weighted network. The nodes and links in the rich club are highlighted, giving \( E_{\geq r} = 6 \) links and \( w_{\geq r} = 4 + 2 + 2 + 3 + 1 + 2 = 14 \) [Figure 6(a)]. The strongest \( E_{\geq r} = 6 \) links of network are weighted and \( \sum_{r=1}^{E_{\geq r}} w_n^{rank} = 4 + 4 + 4 + 3 + 3 + 3 = 21 \). Thus, the \( \phi^w(r) \) obtained is equal to 14/21 [Figure 6(b)]. Finally, the weighted rich-club effect can be detected by measuring the ratio

\[
\rho^w(r) = \frac{\phi^w(r)}{\phi^w_{null}(r)},
\]

where \( \phi^w_{null}(r) \) refers to the weighted rich-club effect assessed on the proper null model. Weight reshuffle is used to produce the null model in this paper; thus the topology of the null model remains intact, and weights are randomly and globally redistributed over the links of the network [12]. When \( \rho^w(r) > 1 \), the original network has a positive weighted rich-club effect, with rich nodes being intertwined with other rich nodes more tightly than randomly expected. In contrast, if \( \rho^w(r) < 1 \), the links among rich nodes are weaker than expected from randomness. In this case, the richness parameter \( r \) is defined as the strength of nodes in order to examine whether the active nodes control the exchange of passenger flows. As it can be seen in Figure 7, the weighted rich-club ratio \( \rho^w(r) \) grows remarkably as a function of the strength of stations. Therefore, the RTS network shows a strong weighted rich-club effect. This finding agrees with previous studies that reported the weighted rich-club effect in the worldwide airport network [12]. Active nodes that handle heavy passenger flows preferentially direct their efforts towards one another, and this tendency becomes more pronounced as the strength of nodes in network increases. Connections among hub stations in the RTS network are characterized by large passenger flows.

3.5. Reliability Analysis of the Shanghai RTS Network. The reliability of Shanghai RTS network is investigated in this section. Several topological parameters can be applied to
evaluate the reliability and vulnerability of the RTS network, such as the largest connected cluster, network efficiency, and network size. In this paper, the largest connected cluster (LCC) is adopted to illustrate the performance changes of RTS network \[16, 17\], which indicates the connectivity of network. In the unweighted networks, the largest connected cluster is defined as follows:

\[
\text{LCC} = \frac{N}{N_0},
\]

where \(N\) is the number of nodes on the largest connected subgraph after attacks and \(N_0\) is the number of nodes on largest connected graph of the initial network. In order to assess the reliability and robustness of the weighted networks, the strength \(s\) is integrated with the largest connected cluster for the weighted case and \(\text{LCC}^w\) is defined by the following equation:

\[
\text{LCC}^w = \frac{S}{S_0},
\]

where \(S\) is the sum of strength of nodes on the largest connected subgraph after attacks and \(S_0\) is the sum of strength of nodes on largest connected graph of the initial network.

It is generally known that when the RTS network is attacked by intentional attacks or random attacks, this will result in failure of some nodes and their connections, and the network might be compromised and even breaks down. If the failure ratio of nodes exceeds the critical threshold, the network will disintegrate into smaller subnetworks and become disconnected. Cohen et al. \[13, 18\] discussed the critical fraction \(p_c\) of the Internet when subjected to random removal and intentional removal, respectively. Zhang et al. \[10\] had done similar researches for the subway network. Combining previous literatures and characteristics of the RTS network failure, the critical threshold \(p_c\) of the fraction of removed nodes is defined as follows. When the nodes are removed from the network one by one and the removed fraction \(p < p_c\), the network has troubles but cannot collapse; otherwise the network almost breaks down when \(p \geq p_c\): in this case, the largest connected cluster for the weighted network is less than 0.05. In this section, the weighted largest connected cluster \(\text{LCC}^w\) is considered as a function of friction of removed nodes \(p\).

In this paper, the critical fraction \(p_c\) of Shanghai RTS network is discussed and compared to random weighted network produced by weight and link reshuffle (i.e., the ties of the network, with their attached weights, are reshuffled in a way that the degree distribution \(p(k)\) is preserved \[19\]). Figure 8 presents the performance changes of the largest connected cluster of the Shanghai RTS network under different attack strategies. Here, WR represents random attacks on the RTS network, WI represents intentional attacks on the RTS network, RWR denotes random attacks on the random weighted network, and RWI denotes intentional attacks on the random weighted network. Meanwhile, the intentional attacks indicate that the largest strength nodes are removed from the network one by one. From Figure 8, the performance and critical thresholds of the RTS network under different attack strategies can be obtained. The critical thresholds for different network failures are \(p_c(WR) = 0.5664\), \(p_c(WI) = 0.2203\), \(p_c(RWR) = 0.5210\), and \(p_c(RWI) = 0.2517\). From the above critical thresholds of the fraction of removed nodes of the RTS network, it is obvious that the intentional attacks can cause more damage than random attacks both to the RTS network and to random weighted network. Furthermore, the intentional attacks are more detrimental to the RTS network than to the random weighted network, but the random attacks can cause slightly more damage to the random weighted network than to the RTS network.

4. Conclusions and Future Research

In this paper, Shanghai RTS network of China is investigated by using complex network theory. The topological properties
and the connectivity, robustness, and reliability of the RTS network are studied from a complex weighted networks perspective.

Simulation results show that the behavior of strength follows a power-law distribution and exhibits heavy tailed behavior; that is, the network possesses hub nodes with high traffic. Moreover, the dynamical properties of a weighted network may differ significantly from its topological properties. Although many nodes possess the same degree, the traffic handled by each rail station may differ significantly.

The weighted clustering coefficient is a measure of the local cohesiveness and provides global information on the correlation between weights and topology by comparing them with their topological clustering. As \( C_w \) is larger than \( C \), it can be concluded that triangles are more likely constructed by edges with larger weights. However, compared with other international metropolises, the local connectivity of the RTS network is very poor. Along with the weighted clustering coefficient, the weighted average nearest neighbors degree is also introduced in this paper and this study finds that the heavily weighted edges are linked to the large-degree neighbors.

Consistently with previous research in worldwide airport network, the RTS network has a positive weighted rich-club effect, indicating that active nodes handling heavy passenger flows are more likely to direct their efforts towards one another. Furthermore, this tendency becomes more pronounced as the strength of nodes in network increases.

The reliability of Shanghai RTS network is investigated in the last section and the findings disclose that the damage caused by intentional attacks is larger than random attacks both to the RTS network and random weighted network. Moreover, the intentional attacks can cause more damage to the RTS network than to the random weighted network, but the reliability of Shanghai RTS network is slightly higher than the random weighted network subjecting to the random attacks.

Complex weighted analysis is a powerful tool for understanding complex architecture of real weighted networks. Indeed, the analysis of the weighted quantities and network reliability provide possibilities for risk analysis and policy decisions through learning structural organization of the network. From the previous literature and research, it is known that there are many other attack strategies and topological parameters for evaluating the reliability and robustness of networks. It would be interesting to study the performance changes of the weighted network under different attack strategies. Another possible extension would be to conduct a comparison analysis of topological and dynamic properties for the weekend and weekday RTS networks and this study may get some interesting results.

Competing Interests

The authors declare that there are no competing interests regarding the publication of this manuscript.

Acknowledgments

The authors would like to thank the Shanghai Shentong Metro Company, for providing valuable information about the Shanghai RTS network for academic and research activities.

References


